

This document contains the **post-print pdf-version** of the refereed paper:

“Multi-purpose economic optimal experiment design applied to model based optimal control”

by *Dries Telen, Boris Houska, Filip Logist and Jan Van Impe*

which has been archived on the university repository Lirias (<https://lirias.kuleuven.be/>) of the KU Leuven.

The content is identical to the content of the published paper, but without the final typesetting by the publisher.

When referring to this work, please cite the full bibliographic info:

D. Telen, B. Houska, F. Logist, J. Van Impe (2016). Multi-purpose economic optimal experiment design applied to model based optimal control, Computers & Chemical Engineering, 94, 212-220.

The journal and the original published paper can be found at:

<http://www.journals.elsevier.com/computers-and-chemical-engineering>

<http://www.sciencedirect.com/science/article/pii/S0098135416302241>

The corresponding author can be contacted for additional info.

Conditions for open access are available at:

<http://www.sherpa.ac.uk/romeo/>

Multi-purpose economic optimal experiment design applied to model based optimal control

D. Telen^a, B. Houska^b, F. Logist^a, J. Van Impe^{a,*}

^a*KU Leuven, Chemical Engineering Department, BioTeC+ & OPTEC,
Gebroeders De Smetstraat 1, 9000 Gent, Belgium*

^b*ShanghaiTech University, School of Information Science and Technology,
319 Yueyang Road, Shanghai 200031, China*

Abstract

In contrast to classical experiment design methods, often based on alphabetic criteria, economic optimal experiment design assumes that our ultimate goal is to solve an optimization or optimal control problem. As the system parameters of physical models are in practice always estimated from measurements, they cannot be assumed to be exact. Thus, if we solve the model based optimization problem using the estimated, non-exact parameters, an inevitable loss of optimality is faced. The aim of economic optimal experiment design is precisely to plan an experiment in such a way that the expected loss of optimality in the optimization is minimized. This paper analyzes the question how to design economic experiments under the assumption that we have more than one candidate objective function. Here, we want to take measurements and estimate the parameters before we actually decide which objective we want to minimize.

Keywords: Optimal experiment design, optimality loss, multi-purpose design, optimal control, variance-covariance matrix

1. Introduction

Model-based optimization is a mathematical tool that has applications in almost all fields of engineering. However, whenever model-based optimization is used to simulate or control real-world processes, an application of computer based numerical methods has to be preceded by experiments that allow us to identify a suitable model and to estimate its associated parameters. Optimal experiment design methods, as originally invented by Fisher (Fisher, 1935), are sought to employ optimization methods already before or during an actual experiment is performed in order to design the experiment in such a way that its expected information content is maximized.

Historically, optimal experiment design (OED) methods have been developed by many authors, for regression models one of the earliest works is Kiefer and Wolfowitz (1959) while the field for nonlinear dynamic systems started with Espie and Macchietto (1989). For a recent discussion on the state-of-the-art the reader is referred to Franceschini and Macchietto (2008) for an overview. The question how to formulate the objective of OED mathematically has no universal answer and from this perspective it is not surprising that many suggestions have been made in the literature. Existing approaches are usually based on minimizing a “suitable” scalar

*Corresponding author

Email address: jan.vanimpe@cit.kuleuven.be (J. Van Impe)

measure of the inverse of the Fisher information matrix (or a direct approximation of the parameter variance-covariance matrix (Heine et al., 2008)) such as the trace, determinant, maximum eigenvalue, or maximum diagonal element leading to the so-called A-, D-, E-, or M-criterion, which also have statistical interpretations (Franceschini and Macchietto, 2008). Here, an empirical observation (see, e.g., Telen et al. (2013)) is that for models with low complexity and a few parameters only, it is—at least from a practical perspective—not excessively important which of the above mentioned objectives is chosen as they often lead to very similar inputs. This empirical observation might also be expected intuitively, since all these design criteria aim at maximizing the “*information content*” of the experiment in one or the other sense.

However, once we consider more complex models with a moderate to large amount of unknown parameters, different OED objectives may lead to significantly different experiments and, in this case, the construction of the OED objective does itself become a modeling problem. In this context, it is important to be aware of the fact that many famous experiment design criteria such as the A- and E-criterion are not even invariant under affine transformations of the parameters (Franceschini and Macchietto, 2008). Consequently, it might not even be clear how to choose a proper scaling of the objective. One way to deal with this issue is to solve a multi-objective OED problem, where a large number of OED problems with different candidate objectives is solved. In this case, the decision which of the corresponding experiments will be implemented, is left to a human decision-maker (Telen et al., 2012). It is clear that such a multi-objective approach is an effective solution for problems, where it is enough to take a small amount of candidate objectives into account. However, as soon as we start screening a 3-, 4- or even higher dimensional Pareto front, the multi-objective approach is limited by the curse of dimensionality.

This paper is about how we can construct suitable objectives for optimal experiment design under the assumption that our ultimate goal is to solve model-based optimization or optimal control problems based on the parameters that are found from the experiment. In this context, it is interesting to mention that for linear systems it is a well-established concept to design experiments with respect to the intended model application (Gevers and Ljung, 1986). This concept has also been elaborated in the context of joint design for control and identification (Hjalmarsson, 2009; Larsson et al., 2015). In Hjalmarsson (2009) this idea has been elaborated and generalized for a broad application spectrum by introducing a generic concept considering quadratic performance degradation costs that can be used to quantify the goal of the experiment design, while in Larsson et al. (2015) the methodology is applied to an industrial case study. Moreover, in Recker et al. (2012) the intended use of the model is taken heuristically into account for the first time for formulating the objectives in nonlinear model-based optimization and optimal experiment design problems. The problem formulation proposed in Houska et al. (2015) leads to a recent concept named “*Economic Optimal Experiment Design*”, which is reviewed in Section 2. The main contribution of this paper is that we extend the idea of economic optimal experiment design for the case that we have more than one application in mind in Section 3. In other words, we want to design experiments that allow us to estimate parameters before we choose an objective function that we want to minimize based on the estimated parameters. Here, our assumption is that we have a set of candidate model applications, i.e., objective functions, in mind when we design the experiment. This can be a typical problem in large reaction networks or plant wide dynamic models. Focussing on the production of some specific desired products (or different cell growth aims), will require an accurate estimation of the (kinetic) parameters in the corresponding reaction paths/sub units. This can lead

to correlated requirements if the reaction paths/sub units are similar or overlapping. If these are hardly overlapping, a correlation between the different economic objectives will not be expected. The corresponding mathematical problem formulation leads to a non-convex min-max optimization problem, which can be reformulated in the form of an equivalent standard nonlinear programming problem, as discussed in Theorem 3.1 of this paper. Section 4 introduces the two illustrative case studies, namely, the Droop model and the Lee-Ramirez bioreactor model of increasing complexity. The numerical results of the case study are described in Section 5. Section 6 concludes the paper.

Notation.

Besides mathematical standard notation, we denote with \mathbb{S}_{++}^n the set of symmetric positive definite matrices. Additionally, we write $A \preceq B$ for two symmetric matrices A and B if the matrix $B - A$ is positive semi-definite. The notation A^\dagger denotes the Moore-Penrose pseudo-inverse of the matrix A .

2. Economic optimal experiment design

In this section, we review the main idea of *economic optimal experiment design* by briefly summarizing some of the basic concepts that have originally been proposed in Houska et al. (2015). For this aim, we start with the most simple case that we are interested in an unconstrained least-squares parameter estimation problem of the form:

$$\min_p \quad \frac{1}{2} \|H(u, p) - \eta\|_{\Sigma^{-1}}^2 + \frac{1}{2} \|p - \hat{p}\|_{\Sigma_0^{-1}}^2. \quad (1)$$

Here, $H : \mathbb{R}^{n_u} \times \mathbb{R}^{n_p} \rightarrow \mathbb{R}^m$ denotes a given measurement function and $\Sigma \in \mathbb{S}_{++}^m$ the variance-covariance matrix of the measurement error. Additionally, an initial parameter estimate $\hat{p} \in \mathbb{R}^{n_p}$ is taken into account, whose precision is assumed to be given in the form of the variance-covariance matrix $\Sigma_0 \in \mathbb{S}_{++}^{n_p}$.

The aim of optimal experiment design is to find an optimal input $u \in \mathbb{R}^{n_u}$, which maximizes the information content of an experiment. Here, information is often quantified in terms of the so-called Fisher information matrix:

$$\mathcal{F}(u, p) := \Sigma_0^{-1} + \left(\frac{\partial H(u, p)}{\partial p} \right)^\top \Sigma^{-1} \frac{\partial H(u, p)}{\partial p},$$

whose inverse $V(u, p) := \mathcal{F}(u, p)^{-1}$ can be interpreted as an affine approximation of the variance-covariance matrix of the predicted parameter estimate (Ljung, 1999; Pukelsheim, 1993). The standard identification procedure based optimal experiment design consists of the following steps:

- 1) Choose a scalar experiment design criterion $\Phi : \mathbb{S}_{++}^{n_p} \rightarrow \mathbb{R}$ and solve the input design problem:

$$u^* \in \underset{u}{\operatorname{argmin}} \Phi(V(u, \hat{p})) \quad \text{subject to} \quad G(u) \leq 0$$

at the best available parameter estimate \hat{p} . Here, $G : \mathbb{R}^{n_u} \rightarrow \mathbb{R}^{n_G}$ is an inequality constraint function modeling the domain of realizable inputs.

- 2) Implement the control u^* and collect measurements.
- 3) Solve the parameter estimation problem (1) and store the new parameter estimate p^* .

- 4) Stop if $\Phi(V(u^*, p^*)) < \text{TOL}$ for a desired accuracy tolerance TOL.
- 5) Set $\hat{p} \leftarrow p^*$ and $\Sigma_0 \leftarrow V(u^*, p^*)$ and continue with Step 1.

Clearly, in the above outlined optimal experiment design procedure, the choice of the scalar design criterion Φ can have a large influence on how the above identification procedure performs. Examples for traditional designs are the A-criterion, E-criterion, and D-criterion, which aim at minimizing the trace, maximum eigenvalue, or determinant of the variance-covariance matrix, respectively. However, these choices are rather ambiguous and, in particular, in the literature on traditional optimal experiment design approaches there is often no advice on how to systematically refine the design criterion if the above identification loop is repeated more than once. This is in contrast to economic optimal experiment (Houska et al., 2015). Here, the underlying assumption is—in the easiest case—that our ultimate goal is to solve an optimization problem of the form:

$$u^*(p) \in \underset{u}{\operatorname{argmin}} F(u, p) \quad \text{subject to} \quad G(u) \leq 0, \quad (2)$$

whose objective function $F(\cdot, p) : \mathbb{R}^{n_u} \rightarrow \mathbb{R}$ depends on the unknown parameter p . The inequality constraint function $G : \mathbb{R}^{n_u} \rightarrow \mathbb{R}^{n_G}$ is assumed to be independent of p . Clearly, if we solve Problem (2) based on a parameter estimate instead of the exact but unknown value, we will in general obtain an optimality gap. Mathematically, this optimality gap can be defined as:

$$\Delta(p) := F(u^*(p), p_{\text{nature}}) - F(u^*(p_{\text{nature}}), p_{\text{nature}}),$$

where p_{nature} denotes the exact but unknown parameter. Now, the aim of the identification procedure is to determine the parameter p in such a way that the expected loss of optimality, i.e., $\mathbb{E}_p \{\Delta(p)\}$ is minimized.

Unfortunately, the exact expectation value $\mathbb{E}_p \{\Delta(p)\}$ is in general rather difficult to compute, since the evaluation of the function Δ requires us to solve a parametric nonlinear programming problem. However, under the additional assumption that the functions F and G are twice continuously differentiable and that $u^*(p_{\text{nature}})$ is a regular minimizer, the function Δ can be approximated by a second-order Taylor expansion. In order to elaborate this expansion, we denote the multiplier of the inequality constraints in Problem (2) with $\mu^*(p) \in \mathbb{R}^{n_G}$. Now, it follows from the complementarity condition that the optimality gap can equivalently be written in the form (Houska et al., 2015):

$$\Delta(p) = L(r^*(p), p_{\text{nature}}) - L(r^*(p_{\text{nature}}), p_{\text{nature}}),$$

where we have introduced the Lagrangian function:

$$L(r, p) = F(u, p) - \mu^\top G(u),$$

as well as the short-hand:

$$r^*(p) := (u^*(p)^\top, \mu^*(p)^\top)^\top,$$

denoting the primal-dual solution of Problem (2) in dependence on p . We can exploit the first order stationarity condition:

$$\frac{\partial L(r^*(p), p)}{\partial u} = 0,$$

which must be satisfied for all p in a sufficiently small neighborhood of p_{nature} recalling that $u^*(p_{\text{nature}})$ is assumed to be a regular KKT point (Nocedal and Wright, 2006). This implies in particular that we have:

$$\Delta(p_{\text{nature}}) = 0 \quad \text{as well as} \quad \frac{\partial \Delta}{\partial p}(p_{\text{nature}}) = 0.$$

Next, the second order derivative of the function Δ can be written in the form:

$$\frac{\partial^2 \Delta}{\partial p^2}(p_{\text{nature}}) = W(p_{\text{nature}}),$$

with:

$$W(p) = J(p)^\top L_{rr}(r^*(p), p) J(p),$$

where we use the shorthands:

$$J = \frac{\partial}{\partial p} r^*, \quad L_{rr} = \frac{\partial^2}{\partial r^2} L, \quad \text{and} \quad L_{rp} = L_{pr}^\top = \frac{\partial^2}{\partial r \partial p} L.$$

Notice that the above expression can be simplified further, since $u^*(p_{\text{nature}})$ is assumed to be a regular KKT point such that it follows from the implicit function theorem for parametric optimization problems (Robinson, 1980) that we have:

$$J(p) = -(L_{rr}(r^*(p), p))^\dagger L_{rp}(r^*(p), p).$$

Using this relation, we can find an explicit expression for the function W , which is given by:

$$W(p) = L_{pr}(r^*(p), p) (L_{rr}(r^*(p), p))^\dagger L_{rp}(r^*(p), p). \quad (3)$$

Consequently, we can approximate the expected loss of optimality by its second order Taylor expansion, given by:

$$\begin{aligned} \mathbb{E}_p \Delta(p) &\approx \frac{1}{2} \mathbb{E}_p \left((p - p_{\text{nature}})^\top W(p_{\text{nature}}) (p - p_{\text{nature}}) \right) \\ &= \frac{1}{2} \text{Trace} \left(W(p_{\text{nature}}) \mathbb{E}_p \{ (p - p_{\text{nature}}) (p - p_{\text{nature}})^\top \} \right). \end{aligned}$$

Motivated by this expression for the second order expansion of the expected loss of optimality we introduce the weighted A-criterion of the form:

$$\forall \Sigma_p \in \mathbb{S}_+^{n_p}, \quad \Phi_{\text{Economic}}(\Sigma_p) := \frac{1}{2} \text{Trace}(W(\hat{p}) \Sigma_p).$$

Here, the second order matrix W is evaluated at the currently best available estimate \hat{p} , since the exact parameter p_{nature} is unknown.

Remark 2.1. A drawback of optimal experiment design based on Fisher matrices (or equivalently linear approximations of the variance-covariance matrix) is that for nonlinear systems the initial parameter estimate \hat{p} has to be used as a linearization point, although \hat{p} may be a very inaccurate guess of the true parameter value p_{nature} . This drawback is not resolved by economic experiment design: in the above formula not only the variance-covariance matrix $V(u, \hat{p})$ uses \hat{p} as a linearization point, but also the weighting matrix $W(\hat{p})$ is evaluated at \hat{p} , as p_{nature} is unknown. One possible way to reduce the error that is associated by this approximation is to use robust optimal experiment design K rkel et al. (2004); Goodwin et al. (2007); Ostrovsky et al. (2013); Li et al. (2008); Galvanin et al. (2010); Telen et al. (2014); Mesbah

and Streif (2015). In principle, the methods from robust OED could be applied to robustify the evaluation of the weighting matrices W as well as the evaluation of the variance matrix V against inaccurate initial parameters guesses. However, such a robust economic OED approach is beyond the scope of this article. Nevertheless, Section 3 discusses an approach that "robustifies" economic OED with respect to the choice of the objective (yet not with respect to inaccurate initial parameter guesses).

Based on this definition, the economic optimal experiment design proceeds in exactly the same way as traditional economic optimal experiment design formulations with the only difference that a very particular choice for the scalar design criterion, namely the function Φ_{Economic} , is used solving the optimization problem:

$$u^* \in \underset{u}{\operatorname{argmin}} \Phi_{\text{Economic}}(V(u, \hat{p})) \quad \text{subject to} \quad G(u) \leq 0 \quad (4)$$

in Step 1) of the above outlined identification loop. Here, it should be noticed that the weighting matrix $W(\hat{p})$ is re-evaluated everytime a new estimate \hat{p} is available, i.e., the design criterion is refined during the procedure. Notice that we first proposed this economic optimal experiment design procedure in Houska et al. (2015), where also the more general case that additional constraints on u are present in the optimal experiment design problem (4) and nominal optimization problem (2). In this context, it is also worth noting that the design criterion Φ_{Economic} is—similar to the D-criterion but in contrast to the standard A- and E-criterion—invariant under affine scaling transformations of the parameter vector.

Remark 2.2. Notice that the notation in this paper is different from Houska et al. (2015), where an additional dynamic system constraints is taken into account explicitly. However, by using a single-shooting discretization (Leineweber et al., 2003) optimal control problems can always be reformulated as a standard nonlinear optimization problem such that the notation can be simplified, as suggested in the current manuscript. Thus, the weighting matrix $W(p)$ can be computed from Equation (3). At this point mixed second order derivatives of the Lagrangian function L are needed, which can be computed approximately by using finite differences or more accurately by using automatic differentiation tools as Houska et al. (2011); Andersson et al. (2012).

3. Multi-purpose economic optimal experiment design

In practical applications, we may want to take one experiment for estimating parameters, but later use these identified parameters for various purposes. Here, "various purposes" means that we do not have one single scalar objective function F , but more than one candidate objective functions F_1, \dots, F_m . The optimization problem of our interest can in this case be written in the form:

$$\min_u \sum_{i=1}^m \alpha_i F_i(u, p) \quad \text{s.t.} \quad G(u) \leq 0. \quad (5)$$

Here, $\alpha \in S$ denotes a weighting vector, which is assumed to be in the unit simplex:

$$S = \left\{ \alpha \in \mathbb{R}^m \left| \sum_{i=1}^m \alpha_i = 1, \forall i \in \{1, \dots, m\}, \alpha_i \geq 0 \right. \right\}.$$

For example, if we choose $\alpha = e_1 = (1, 0, \dots, 0)^\top$ this would mean that our goal is to minimize the function F_1 . If we choose for α the i -th unit vector e_i , this corresponds to minimizing the i -th objective function. In the most general case, our goal

is to optimize a linear combination of the given candidate objectives. However, the main problem that we are facing in this context is that the weighting vector $\alpha \in S$ is in our setting not known at the time when we design the experiment. That is, we want to estimate the parameter p first in view of a model based optimization but choose α in the objective function later.

In order to apply the analysis from the previous section to this situation, we introduce the parametric Lagrangian function:

$$L(\alpha, r, p) = \sum_{i=1}^m \alpha_i F_i(u, p) - \mu^\top G(u), \quad (6)$$

where $r = (u, \mu)$ is a stacked version of the primal-dual solution of the optimization problem (5). The weighting matrix $W(\alpha, p)$ is now defined in complete analogy to our previous considerations:

$$\begin{aligned} W(\alpha, p) \\ = L_{pr}(\alpha, r^*(p), p) (L_{rr}(\alpha, r^*(p), p))^\dagger L_{rp}(\alpha, r^*(p), p), \end{aligned}$$

with the only difference that we have now an additional dependency on the parameter α . Using this notation, the multi-purpose economic optimal experiment design problem can be written in the form:

$$\min_u \max_{\alpha \in S} \frac{1}{2} \text{Trace}(W(\alpha, \hat{p})V(u, \hat{p})) \quad , \quad \text{s.t.} \quad G(u) \leq 0. \quad (7)$$

Here, the inner maximization over the weighting vector $\alpha \in S$ takes into account that we do not know α in advance. Thus, we propose to minimize the worst possible expected loss of optimality. Problem (7) is a bilevel optimization problem, whose inner maximization problem is non-concave and there arises the question how this problem can be solved in practice. One of the key contributions of this paper is that we show that Problem (7) can be reformulated into an equivalent standard minimization problem thereby rendering multi-purpose economic OED problems tractable for standard NLP solvers. This result is summarized in the following theorem.

Theorem 3.1. *Let e_i denote the i -th unit vector in \mathbb{R}^m and the functions F_i and G be twice continuously differentiable. Now, the min-max optimization problem (7) is equivalent to the standard minimization problem:*

$$\min_{u, \gamma} \frac{\gamma}{2} \quad \text{s.t.} \quad \begin{cases} \forall i \in \{1, \dots, m\}, \\ \text{Tr}(W(e_i, \hat{p})V(u, \hat{p})) \leq \gamma, \\ G(u) \leq 0. \end{cases} \quad (8)$$

Proof. The main idea is to exploit the fact that the Lagrangian function $L(\alpha, r, p)$ (Equation (6)) is by construction affine in the variable α . Consequently, the function $W(\alpha, p)$ is a “quadratic-over-linear” function in α (see Boyd and Vandenberghe (2004)) and, thus, we have:

$$\theta W(\alpha_1, p) + (1 - \theta)W(\alpha_2, p) \succeq W(\theta\alpha_1 + (1 - \theta)\alpha_2, p)$$

for all $\alpha_1, \alpha_2 \in S$ and all $\theta \in [0, 1]$. This implies that the objective function of the

min-max optimization problem (7), given by:

$$\frac{1}{2} \text{Tr} (W(\alpha, \hat{p}) V(u, \hat{p})) ,$$

is a convex function in α . As the maximum of a convex function over a simplex is always obtained at a vertex, it is enough to evaluate this function at the unit vectors e_i thereby establishing the equivalence between the optimization problems (7) and (8). \diamond

4. Case studies

Two case studies are investigated in this paper. Section 4.1 studies the Droop model which describes the growth of micro algae. Section 4.2 introduces a second case study, the Lee-Ramirez bioreactor. The latter describes the induced foreign protein production by recombinant bacteria in a fed-batch bioreactor.

4.1. The Droop model

The first illustrative case study employed in this paper is the Droop model (Bernard, 2011). It describes the growth of micro algae in a photobioreactor under constant temperature and illumination conditions. The model equations in the interval $t \in [0, T]$ are given by:

$$\dot{S} = -\rho(S)X - D(S - S_{\text{in}}) , \quad (9)$$

$$\dot{Q} = \rho(S) - \mu(Q)Q , \quad (10)$$

$$\dot{X} = \mu(Q)X - DX . \quad (11)$$

Here, the states, S , Q , and X denote the substrate concentration (mg N/L), the intracellular quota (mg N/ mg C), and the biomass concentration (mg C/L). All states are assumed to be measurable with the following measurement variances, $\sigma_S^2 = 1.0 \text{ (mg N/L)}^2$, $\sigma_Q^2 = 1.0 \cdot 10^{-5} \text{ (mg N/ mg C)}^2$, and $\sigma_X^2 = 1.0 \cdot 10^3 \text{ (mg C/L)}^2$, the non-diagonal elements are assumed to be zero. This indicates that in particular the biomass concentration is hard to measure online. The control action is the dilution rate D , while S_{in} is the pre-set substrate concentration in the feed. For all dynamic optimizations, a single shooting approach is employed, where the control action is discretized in 14 steps (each corresponding to a single day), so $u = (D(0), \dots, D(T-1))^T$. The uptake rate is given by the following equation:

$$\rho(S) = \rho_m \frac{S}{S + K_s} , \quad (12)$$

while the growth rate is described by:

$$\mu(Q) = \mu_m \left(1 - \frac{Q_0}{Q} \right) . \quad (13)$$

For this model two objective functions are considered, namely, maximizing the biomass concentration $F_1(u, p) = -X(T)$, and tracking the biomass concentration at 100 mg C/L, $F_2(u, p) = \int_0^T (X(t) - 100)^2 dt$. In the model the following three parameters are of interest for the optimal experiment design procedure, i.e., $p = (\mu_m, K_s, \rho_m)^T$. The total duration of the experiment T is fixed to 14 days. The initial variance in the states required for the variance-covariance matrix computation is considered to be ten times the corresponding measurement error variance while for the parameters, the following initial variances are assumed, $\sigma_{\mu_m}^2 = 0.256 \text{ day}^{-2}$,

$\sigma_{K_s}^2 = 5.625 \text{ mg N/L}^2$, and $\sigma_{\rho_m}^2 = 0.001 (\text{mg N/ (mg C} \cdot \text{day)})^2$. Bounds on the operating conditions and numerical values for the parameters and the remaining constants are given in Table 1. Due to the fact that there are 3 states and 3 parameters of interest, a total of $6 \times 7/2 = 21$ additional states are required for the variance-covariance matrix computation.

Table 1: Overview of the experimental conditions, the employed parameter and constant values.

Experimental conditions	Parameters and constants
$S(0) \in [0, 15] \text{ mg N/L}$	$\mu_m = 1.645 \text{ day}^{-1}$
$Q(0) = 0.10 \text{ mg N/ mg C}$	$K_s = 7.5 \text{ mg N/L}$
$X(0) \in [0, 40] \text{ mg C/L}$	$\rho_m = 0.1 \text{ mg N/ (mg C} \cdot \text{day)}$
$D \in [0, 0.5] \text{ day}^{-1}$	$Q_0 = 0.04 \text{ mg N/ (mg C)}$
	$S_{in} = 4.0 \text{ mg N/L}$

4.2. The Lee-Ramirez bioreactor

The case study presented in this section is the Lee-Ramirez bioreactor. It describes the induced foreign protein production by recombinant bacteria in fed-batch bioreactors (Lee and Ramirez, 1996). The differential equations in an interval $[0, T]$ are given by:

$$\dot{x}_1 = u_1 + u_2, \quad (14)$$

$$\dot{x}_2 = \mu(x_3, x_5, x_6, x_7)x_2 - (u_1 + u_2)\frac{x_2}{x_1}, \quad (15)$$

$$\dot{x}_3 = C_{s,in}\frac{u_1}{x_1} - (u_1 + u_2)\frac{x_3}{x_1} - \mu(x_3, x_5, x_6, x_7)\frac{x_2}{Y}, \quad (16)$$

$$\dot{x}_4 = \pi(x_3, x_5)x_2 - (u_1 + u_2)\frac{x_4}{x_1}, \quad (17)$$

$$\dot{x}_5 = C_{i,in}\frac{u_2}{x_1} - (u_1 + u_2)\frac{x_5}{x_1}, \quad (18)$$

$$\dot{x}_6 = -k_1(x_5)x_6, \quad (19)$$

$$\dot{x}_7 = k_1(x_5)(1 - x_7), \quad (20)$$

in which there are the specific growth rate μ , the foreign production rate π and shock and recovery parameter k_1 :

$$\mu(x_3, x_5, x_6, x_7) = \frac{\mu_{\max}x_3}{K_{CN} + x_3 + \frac{x_3^2}{K_s}} \left(x_6 + x_7 \frac{K_{CI}}{K_{CI} + x_5} \right), \quad (21)$$

$$\pi(x_3, x_5) = \frac{f_{\max}x_3}{K_{CN} + x_3 + \frac{x_3^2}{K_s}} \frac{f_{Io} + x_5}{K_I + x_5}, \quad (22)$$

$$k_1(x_5) = \frac{k_{11}x_5}{K_{IX} + x_5}. \quad (23)$$

The states are x_1 , the reactor volume (L), x_2 , the cell density (g/L), x_3 , the nutrient concentration (g/L), x_4 , the foreign protein concentration (g/L), x_5 , the inducer concentration (g/L), x_6 the inducer shock factor on cell growth rate (-) and x_7 , the inducer recovery factor on cell growth rate (-). The following states are assumed to be measurable with the following measurement variances $\sigma_{x_1}^2 = 0.01 \text{ (L)}^2$, $\sigma_{x_2}^2 = 0.01 \text{ (g/L)}^2$, $\sigma_{x_3}^2 = 0.01 \text{ (g/L)}^2$, $\sigma_{x_4}^2 = 0.01 \text{ (g/L)}^2$. The remaining non-diagonal elements are assumed to be zero. The initial variance in the states required for the variance-covariance matrix computation is considered to be ten percent of the

Table 2: Overview of the experimental conditions, the employed parameter and constant values.

Parameters and constants	Experimental conditions
$\mu_{\max} = 1.0$ (1/h)	$C_{s,\text{in}} = 100$ (g/L)
$K_{C_N} = 14.35$ (g/L)	$C_{i,\text{in}} = 4.0$ (g/L)
$K_s = 111.5$ (L/g)	$u_1 \in [0, 1]$ (L/h)
$K_{C_I} = 0.22$ (g/L)	$u_2 \in [0, 1]$ (L/h)
$f_{\max} = 0.223$ (1/h)	$x_1(0) = 1$ (L)
$f_{I_0} = 0.0005$ (g/L)	$x_2(0) = 0.1$ (g/L)
$K_I = 0.022$ (g/L)	$x_3(0) = 40$ (g/L)
$k_{11} = 0.09$ (1/h)	$x_4(0) = 0$ (g/L)
$K_{IX} = 0.034$ (g/L)	$x_5(0) = 0$ (g/L)
$Y = 0.51$ (-)	$x_6(0) = 1$ (-)
	$x_7(0) = 0$ (-)

initial value, for the parameters, the initial standard deviations are assumed to be ten percent of the current values. The control actions for this case study are denoted by u_1 , i.e., the glucose feed rate and u_2 , i.e., the inducer feed rate (L/h). The constants $C_{s,\text{in}}$ and $C_{i,\text{in}}$ are pre-set concentration for the inducer and glucose concentration in the feed rate. The final time is fixed $T = 10$ h. For all dynamic optimizations, a single shooting approach is employed, where the controls are discretized in 10 steps (each corresponding to 1 h) $u = (u_1(0), \dots, u_1(T-1), u_2(0), \dots, u_2(T-1))^T$. Note that due to the fact that there are 7 states and 10 parameters of interest, the size of the optimal experiment design problem grows considerably compared with the first case study, $17 \times 18/2 = 153$ additional states for the variance-covariance matrix need to be computed due to the exploitation of symmetry.

The objectives of interest for this case study are the following: maximizing the protein production, $F_1(u, p) = -x_1(T)x_4(T)$, tracking the cell density at 1.0 g/L, $F_2(u, p) = \int_0^T (x_2(t) - 1.0)^2 dt$, and tracking the protein concentration at 1.0 g/L, $F_3(u, p) = \int_0^T (x_4(t) - 1.0)^2 dt$.

5. Simulation results

Both case studies are simulated using the ACADO toolkit (Houska et al., 2011). We first optimized both systems with the parameter values from Table 1 and Table 2 for their respective objective functions. The obtained control profiles allow us to compute the weighting matrices $W(e_i, \hat{p})$, required in each of the multi-purpose optimal experiment design procedures.

5.1. The Droop model

For comparison, we also design an experiment with the traditional A-criterion, i.e., the minimization of the trace of the parameter variance-covariance matrix. The parameter variance-covariance matrix is computed using the approach of (Telen et al., 2013). The employed KKT-tolerance is set to 10^{-5} . The obtained state profiles for both the A-optimal experiment design and the multi-purpose economic experiment design is depicted in Figure 1. Both designs have the same starting substrate concentration, while the biomass concentration in the multi-purpose design starts slightly higher than the A-optimal design. The main difference is in

the resulting control action. The dilution rate reaches its maximum value in the multi-purpose experiment after two days while this is three days in the A-optimal design. Furthermore, the A-optimal design starts first with reducing the dilution rate to zero after 9 days while the multi-purpose design reduces the dilution rate one day later.

The different effect of the two experiments is also visible in the evolution of the predicted parameter variances as illustrated in Figure 2. The A-optimal experiment minimizes the trace of the variance-covariance matrix and thus aims at reducing the largest parameter variance, i.e., reducing the variance of K_s . If we compare this with the evolution for the other two parameters, we observe that the multi-purpose experiment aims at reducing the variance in the other two parameters and results in a higher variance for K_s than in the A-optimal design. So, based on the designed experiments, we see that K_s is the focus of the A-optimal design, while μ_m and ρ_m are more the focus of the multi-purpose design. When we evaluate the multi-purpose objective function for the two designs, we obtain 158 (mg C/L)^2 for the multi-purpose design and 505 (mg C/L)^2 for the A-optimal design. When we investigate to which of the two objective functions this optimality gap relates, we observe that this relates to the tracking objective $F_2 = 1571 \text{ (mg C/L)}^2$. The relative expected performance loss is thus 0.10 for the multi-purpose design versus 0.32 for the A-optimal design.

In order to validate the predicted expected optimality gap, Monte Carlo simulations are performed. The following approach is pursued. For both designs 100 noise realizations (based on the assumed measurement noise from Section 4) are drawn and added to the three state evolutions. Subsequently, we perform a parameter estimation procedure based on these noisy profiles resulting in 100 parameter sets for each design. Based on these parameter sets, 100 optimal control problems for F_2 are solved for each design, as F_2 was the most sensitive objective function with respect to the expected optimality gap. The followed approach is also illustrated in Figure 3. The obtained objective function values are compared with $F_2(u^*(p_{\text{nature}}), p_{\text{nature}}) = 1571$. We report the three quartiles for both designs. The multi-purpose design approach results in $Q_1 = 31.7 \text{ (mg C/L)}^2$, $Q_2 = 106.4 \text{ (mg C/L)}^2$ and $Q_3 = 362.0 \text{ (mg C/L)}^2$ with a predicted value of 158 (mg C/L)^2 while the A-optimal design results in $Q_1 = 298.7 \text{ (mg C/L)}^2$, $Q_2 = 667.7 \text{ (mg C/L)}^2$ and $Q_3 = 988.7 \text{ (mg C/L)}^2$ with a predicted value of 505 (mg C/L)^2 . As the data of the expected optimality gap is skewed, we use the median as the approximation of the expected value. We observe that the predicted expected optimality gaps are in agreement with the results of the numerical experiments illustrating the potential of the proposed economic optimal experiment design framework.

The different effect of the two experiments is also clearly visible in the parameter estimates of the Monte Carlo simulations. The resulting mean estimate and standard deviation for each of the two type experiments are given in Table 3. We observe that the confidence region for μ_m and ρ_m is the smallest in the multi-purpose design while the confidence region of K_s is the smallest in the A-optimal design. This effect has been predicted by the computed variance-covariance matrix elements in the optimal experiment designs. So, this illustrates that not all parameters need to be estimated equally accurate for a given economic objective.

Note. In the presented case study the parameters used in the optimal experiment design corresponded to the parameters of the systems for which *noisy* experiments have been performed. In practice however the parameters in the design differ from the true system parameters. To this extent the experiment needs to be informative

Table 3: Overview of the obtained parameter estimates and their corresponding standard deviation (between brackets).

A-optimal design	Multi-purpose design
$\hat{\mu}_m = 1.695$ (0.0286)	$\hat{\mu}_m = 1.641$ (0.0148)
$\hat{K}_s = 7.844$ (0.1800)	$\hat{K}_s = 8.001$ (0.2164)
$\hat{\rho}_m = 0.109$ (0.00245)	$\hat{\rho}_m = 0.106$ (0.00174)

for a broad range of parameter values. This is the field of *robust optimal experiment design* (Goodwin et al., 2007). The approaches presented in the literature, i.e., a worst case approach (Körkel et al., 2004) or an expected value approach (Ostrovsky et al., 2013; Li et al., 2008; Galvanin et al., 2010; Telen et al., 2014; Mesbah and Streif, 2015), possibly with chance constraints, can be extended to make our presented approach more robust.

5.2. The Lee-Ramirez bioreactor

Also for the second case study two experiments are designed: on the one hand an experiment with the traditional A-criterion and on the other hand the multi-purpose experiment design approach. Note that in this case study the difference between several parameters is multiple orders of magnitude, so the relative parameter variance is employed. The parameter variance-covariance matrix is computed using the approach of Telen et al. (2013). The employed KKT-tolerance is set to 10^{-4} . The obtained state profile evolution for both approaches is presented in Figure 4.

We observe two distinct experiments obtained by the two different approaches. The multi-purpose approach has at 8 hours an hour long glucose feed which is completely absent in the multi-purpose experiment. This glucose feed results in an excitation of the system as there is a sharp reaction in the cell density, nutrient and protein concentration. In contrast the A-optimal design feeds inducer at a mild level at the start of the experiment while the multi-purpose experiment has only a very small inducer feeding period. When the expected parameter variance evolutions are investigated in Figure 5, we observe that for both experiments for four parameters (f_{I_0} , K_I , K_{C_I} and K_{I_X}) there is no influence on their parameter variance. Furthermore, we observe that the main focus of the A-optimal experiment is on reducing (slightly) the variance of k and K_s while the multi-purpose has a lower value for μ_{\max} , f_{\max} and Y .

When we evaluate the two designs we observe a the following difference in the predicted relative optimality gap. For the multi-purpose objective function a relative difference of 0.27 is obtained while for the A-optimal design 0.33 is reached. It is observed that this difference is related to the tracking of the biomass objective, i.e., F_2 . So, in contrast with the first case study where the optimality gap could be reduced by a factor 3, we only observe a reduction of 18% for this case study. To validate the difference between the two designed experiments, a Monte Carlo simulation similar to the first case study, see also Figure 3, is performed. For the multi-purpose experiment, the following quartiles are obtained: $Q_1 = 0.058$, $Q_2 = 0.18$ and $Q_3 = 0.34$ while for the A-optimal experiment, the quartiles are $Q_1 = 0.060$, $Q_2 = 0.22$ and $Q_3 = 0.38$. We see that the first quartile is similar for both approaches, however, for Q_2 and Q_3 this difference is in agreement with the predicted difference. We also report the obtained relative means and variances for the two type of experiments. For the multi-purpose experiment this is 0.23 with variance 0.037 while for the A-optimal experiment design 0.29 and 0.054 is obtained.

So, also in the mean and variance a difference in the two experiments is observed.

Also the obtained parameter estimates after the Monte Carlo simulations are discussed. An overview of the obtained numerical values can be found in Table 4. For both experiments we remark that the parameters (f_{Io} , K_I , K_{CI} and K_{IX}) are impossible to estimate accurately as the obtained variances are significantly larger than the predicted variances from Figure 5. We also observe that the multi-purpose experiment leads to a smaller confidence region for μ_{\max} and f_{\max} which also seems to be in agreement with the predictions. For parameter Y , the A-optimal design leads to slightly smaller confidence region, although the difference is small. Also note that K_s and k_{11} are more the target of the A-optimal design, however, only K_s is more accurately estimated in the A-optimal design contrary to the prediction.

Table 4: Overview of the obtained parameter estimates and their corresponding standard deviation (between brackets).

A-optimal design	Multi-purpose design
$\mu_{\max} = 1.05$ (0.13)	$\mu_{\max} = 1.09$ (0.11)
$K_{CN} = 14.98$ (0.25)	$K_{CN} = 14.22$ (0.28)
$K_s = 115.6$ (0.24)	$K_s = 104.8$ (0.28)
$K_{CI} = 0.20$ (0.26)	$K_{CI} = 0.20$ (0.25)
$f_{\max} = 0.23$ (0.11)	$f_{\max} = 0.23$ (0.08)
$f_{Io} = 0.00050$ (0.30)	$f_{Io} = 0.00044$ (0.27)
$K_I = 0.022$ (0.29)	$K_I = 0.025$ (0.27)
$k_{11} = 0.091$ (0.26)	$k_{11} = 0.098$ (0.21)
$K_{IX} = 0.033$ (0.30)	$K_{IX} = 0.030$ (0.27)
$Y = 0.51$ (0.012)	$Y = 0.51$ (0.019)

6. Conclusions

In this paper, we have introduced a novel class of economic objectives for optimal experiment design under the assumption that the estimated parameters shall be used in possibly more than one application. A key contribution of this paper is presented in Theorem 3.1, where we have proven that min-max multi-purpose economic optimal experiment design problems can be reformulated as standard nonlinear programming problems, which makes these problems tractable for solving with standard nonlinear programming problem solvers. We have illustrated the practical applicability of the proposed formulation with two different case studies, i.e., an illustrative model describing the growth of micro algae and a model describing the induced foreign protein production by recombinant bacteria. We observed that the multi-purpose approach leads in the first case study to a reduction of the worst case expected optimality gap by a factor compared with the A-optimal design while for the second case study a reduction of 18% is observed. In addition, these predictions are validated by Monte Carlo simulations which confirm the predicted expected optimality losses.

Acknowledgments

This research was supported by the KU Leuven Research Council: PFV/10/002 Optimization in Engineering Center OPTEC, the Flemish Government: FWO: FWO KAN2013 1.5.189.13, FWO-G.0930.13; the Belgian Federal Science Policy Office:

IUAP P7 (DYSCO, Dynamical systems, control and optimization, 2012-2017); the National Science Foundation China (NSFC), Nr. 61473185, as well as ShanghaiTech University, Grant-Nr. F-0203-14-012.

References

- Andersson, J., Akesson, J., Diehl, M., 2012. CasADi - a symbolic package for automatic differentiation and optimal control. In: Proceedings of the 6th International Conference on Automatic Differentiation.
- Bernard, O., 2011. Hurdles and challenges for modelling and control of microalgae for CO₂ mitigation and biofuel production. *Journal of Process Control* 21 (0), 1378–1389.
- Boyd, S., Vandenberghe, L., 2004. Convex optimization. Cambridge University Press, Cambridge.
- Espie, D., Macchietto, S., 1989. The optimal design of dynamic experiments. *AIChE Journal* 35, 223–229.
- Fisher, R., 1935. The design of experiments. Oliver & Boyd.
- Franceschini, G., Macchietto, S., 2008. Model-based design of experiments for parameter precision: State of the art. *Chemical Engineering Science* 63, 4846–4872.
- Galvanin, F., Barolo, M., Bezzo, F., Macchietto, S., 2010. A backoff strategy for model-based experiment design under parametric uncertainty. *AIChE Journal* 56, 2088–2102.
- Gevers, M., Ljung, L., 1986. Optimal experiment designs with respect to the intended model application. *Automatica* 22, 543–554.
- Goodwin, G., Rojas, C., Welsh, J., Feuer, A., 2007. Robust optimal experiment design for system identification. *Automatica* 43, 993–1008.
- Heine, T., Kawohl, M., King, R., 2008. Derivative-free optimal experimental design. *Chemical Engineering Science* 63, 4873–4880.
- Hjalmarsson, H., 2009. System identification of complex and structured systems. *European Journal of Control* 15, 275 – 310.
- Houska, B., Ferreau, H., Diehl, M., 2011. ACADO Toolkit - an open-source framework for automatic control and dynamic optimization. *Optimal Control Applications and Methods* 32, 298–312.
- Houska, B., Telen, D., Logist, F., Diehl, M., Van Impe, J., 2015. An economic objective for optimal experiment design of nonlinear dynamic processes. *Automatica* 51, 98–103.
- Kiefer, J., Wolfowitz, J., 1959. Optimum designs in regression problems. *Annals of Mathematical Statistics* 30, 271–294.
- Körkel, S., Kostina, E., Bock, H., Schlöder, J., 2004. Numerical methods for optimal control problems in design of robust optimal experiments for nonlinear dynamic processes. *Optimization Methods and Software Journal* 19 (3-4), 327–338.
- Larsson, C., Rojas, C., Bombois, X., Hjalmarsson, H., 2015. Experimental evaluation of model predictive control with excitation (MPC-X) on an industrial depuranizer. *Journal of Process Control* 31, 1 – 16.

- Lee, J., Ramirez, W., 1996. On-line optimal control of induced foreign protein production by recombinant bacteria in fed-batch reactors. *Chemical Engineering Science* 51, 521–534.
- Leineweber, D., Bauer, I., Bock, H., Schlöder, J., 2003. An efficient multiple shooting based reduced SQP strategy for large-scale dynamic process optimization. Part I: theoretical aspects. *Computers and Chemical Engineering* 27, 157–166.
- Li, P., Arellano-Garcia, H., Wozny, G., 2008. Chance constrained programming approach to process optimization under uncertainty. *Computers and Chemical Engineering* 32, 25 – 45.
- Ljung, L., 1999. *System Identification: Theory for the User*. Prentice Hall.
- Mesbah, A., Streif, S., 2015. A probabilistic approach to robust optimal experiment design with chance constraints. In: *Proceedings of the 9th International Symposium on Advanced Control of Chemical Processes (ADCHEM)*. IEEE, pp. 100–105.
- Nocedal, J., Wright, S., 2006. *Numerical optimization*, 2nd Edition. Springer series in operation research and financial engineering. Springer, New York.
- Ostrovsky, G., Ziyatdinov, N., Lapteva, T., 2013. Optimal design of chemical processes with chance constraints. *Computers and Chemical Engineering* 59, 74 – 88.
- Pukelsheim, F., 1993. *Optimal design of Experiments*. John Wiley & Sons, Inc., New York.
- Recker, S., Kühn, P., Diehl, M., Bock, H., 2012. Sigmappoint approach for robust optimization of nonlinear dynamic systems. In: *Proceeding of SIMULTECH 2012*. pp. 199–207.
- Robinson, S. M., 1980. Strongly regular generalized equations. *Mathematics of Operations Research* 5, 43–62.
- Telen, D., Houska, B., Logist, F., Vanderlinden, E., Diehl, M., Van Impe, J., 2013. Optimal experiment design under process noise using Riccati differential equations. *Journal of Process Control* 23, 613–629.
- Telen, D., Logist, F., Vanderlinden, E., Van Impe, J., 2012. Optimal experiment design for dynamic bioprocesses: a multi-objective approach. *Chemical Engineering Science* 78, 82–97.
- Telen, D., Vercammen, D., Logist, F., Van Impe, J., 2014. Robustifying optimal experiment design for nonlinear, dynamic (bio)chemical systems. *Computers and Chemical Engineering* 71, 415–425.

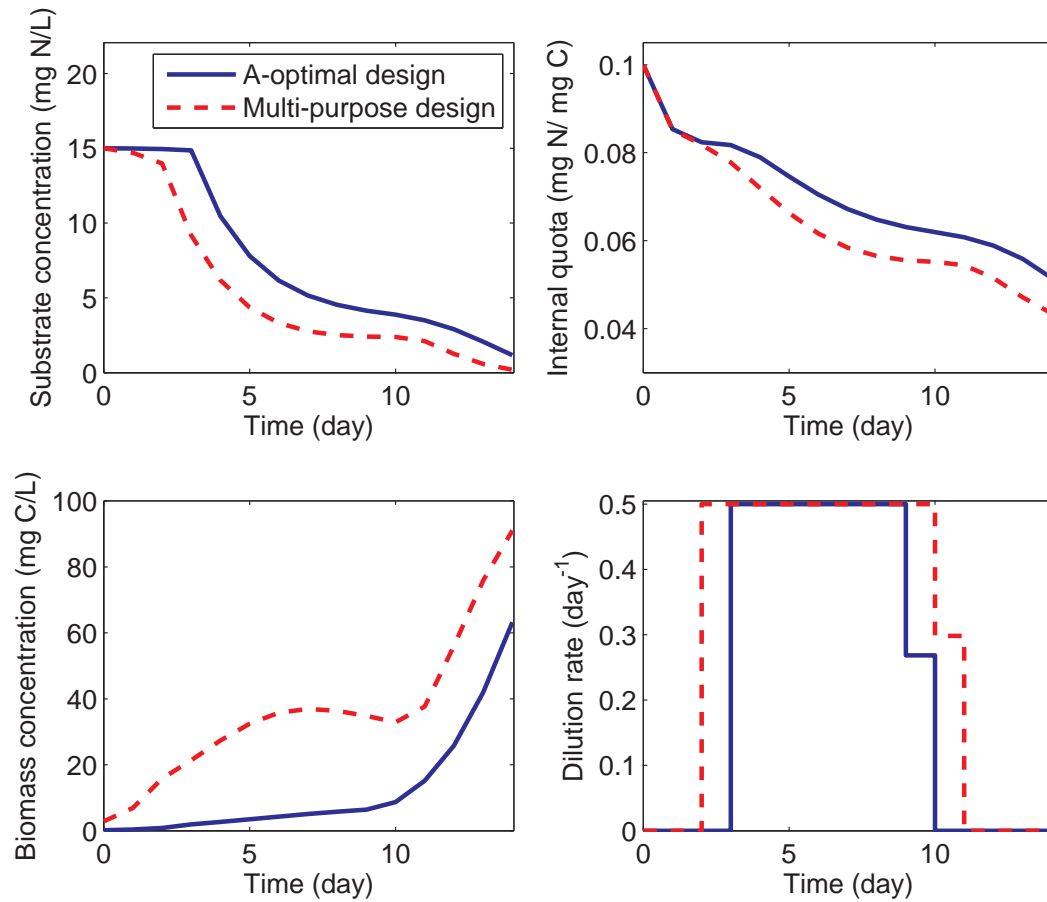


Figure 1: Obtained state profiles and control actions for both the A-optimal and multi-purpose experiment design procedure.

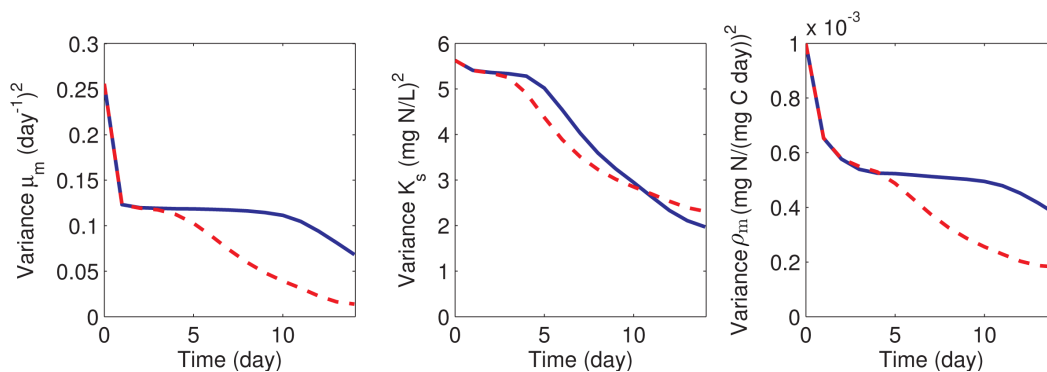


Figure 2: Obtained variance evolution for both the A-optimal and multi-purpose experiment design procedure.

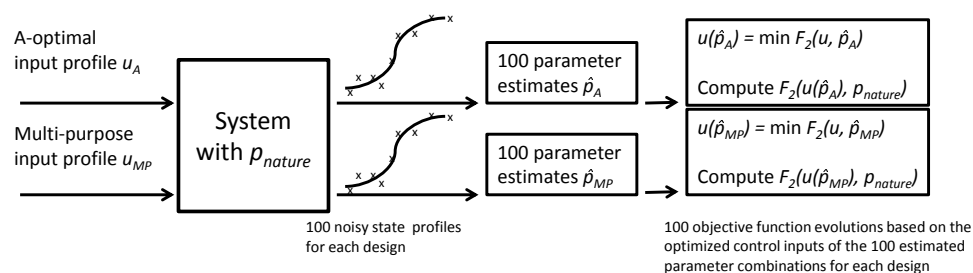


Figure 3: Numerical validation approach for the predicted expected optimality gap for case study 1.

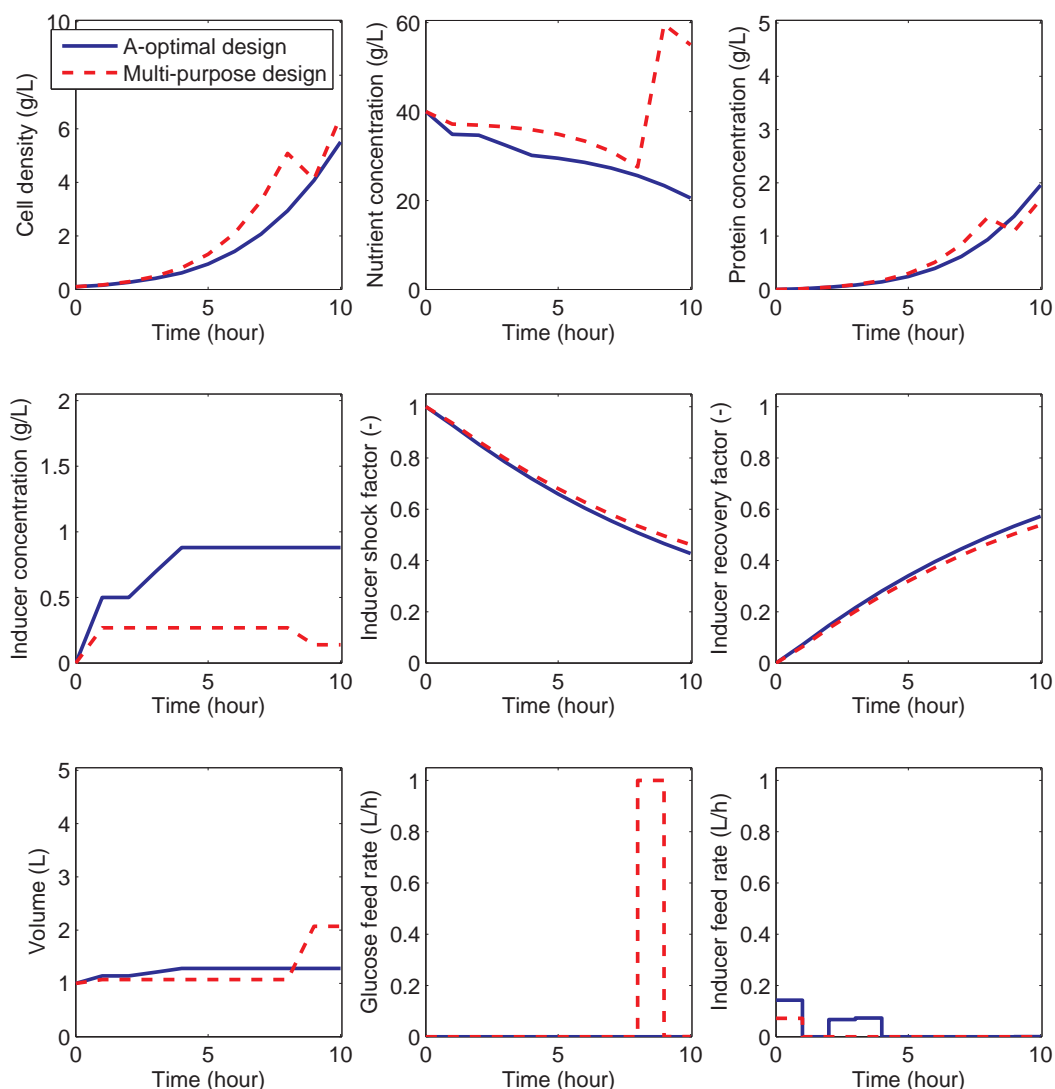


Figure 4: Obtained state profiles and control actions for both the A-optimal and multi-purpose experiment design procedure for the Lee-Ramirez .

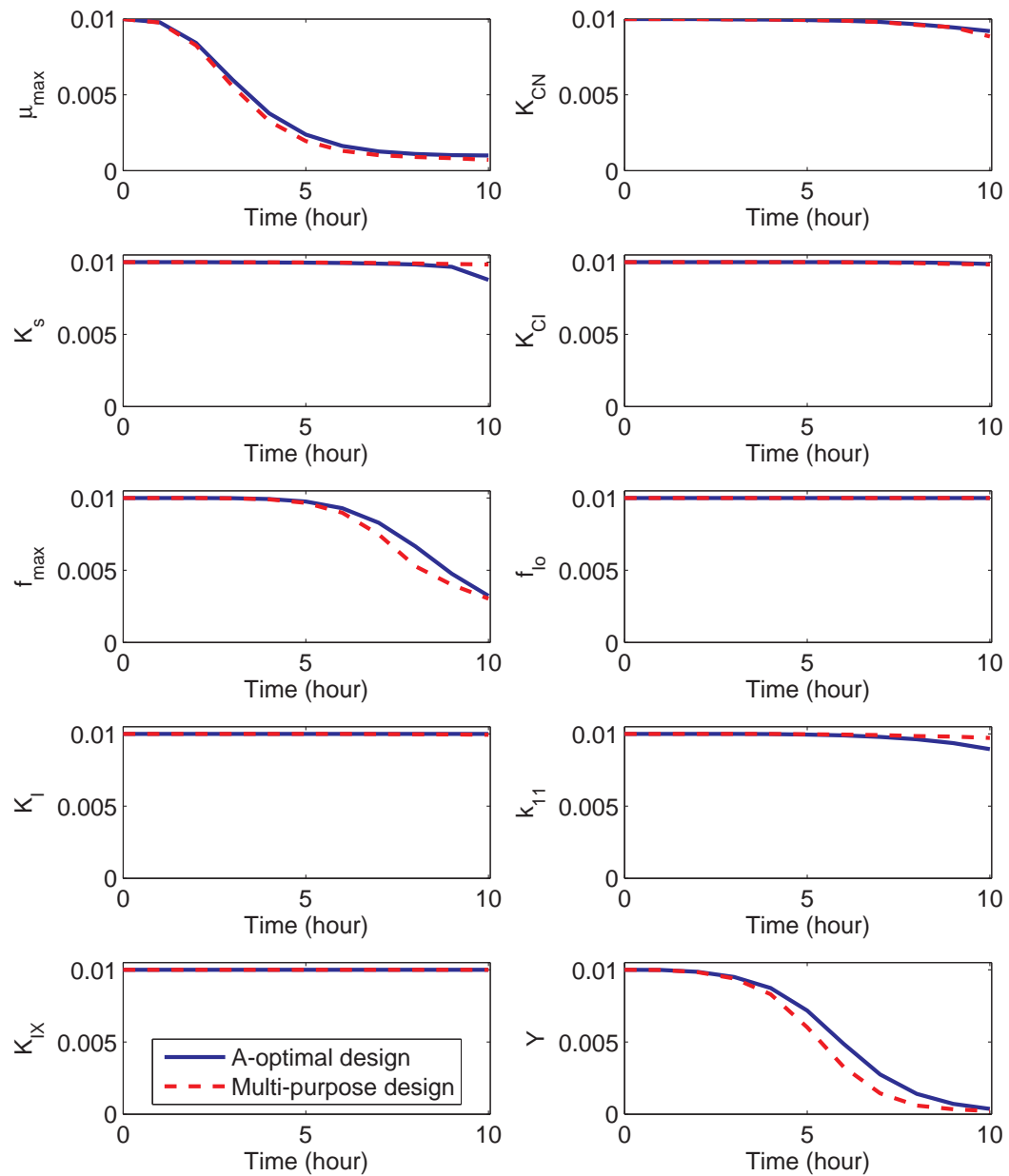


Figure 5: Obtained variance evolution for both the A-optimal and multi-purpose experiment design procedure for the Lee-Ramirez bioreactor.